Genomics 106 (2015) 88–95

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

GAA triplet-repeats cause nucleosome depletion in the human genome

Hongyu Zhao [a,b], Yongqiang Xing [b], Guoqing Liu [b], Ping Chen [c], Xiujuan Zhao [b], Guohong Li [c,*], Lu Cai [b,*]

[a] School of Physical Science and Technology, Inner Mongolia University, Hohhot 010021, China
[b] School of Mathematics, Physics and Biological Engineering, Inner Mongolia University of Science and Technology, Baotou 014010, China
[c] National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

ARTICLE INFO

Article history:
Received 22 February 2015
Received in revised form 15 June 2015
Accepted 18 June 2015
Available online 20 June 2015

Keywords:
Trinucleotide repeats
Friedreich's ataxia
Chromatin structures

ABSTRACT

Although there have been many investigations into how trinucleotide repeats affect nucleosome formation and local chromatin structure, the nucleosome positioning of GAA triplet-repeats in the human genome has remained elusive. In this work, the nucleosome occupancy around GAA triplet-repeats across the human genome was computed statistically. The results showed a nucleosome-depleted region in the vicinity of GAA triplet-repeats in activated and resting $CD4^+$ T cells. Furthermore, the A-tract was frequently adjacent to the upstream region of GAA triplet-repeats and could enhance the depletion surrounding GAA triplet-repeats. *In vitro* chromatin reconstitution assays with GAA-containing plasmids also demonstrated that the inserted GAA triplet-repeats destabilized the ability of recombinant plasmids to assemble nucleosomes. Our results suggested that GAA triplet-repeats have lower affinity to histones and can change local nucleosome positioning. These findings may be helpful for understanding the mechanism of Friedreich's ataxia, which is associated with GAA triplet-repeats at the chromatin level.

1. Introduction

Nucleosome formation on repeated DNA fragments has been investigated for many years [1,2]. The previous works showed that CAG trinucleotide repeats associated with diseases such as myotonic dystrophy type 1 and spinocerebellar ataxia type 1 are strong nucleosome positioning signals [2–5]. The CGG triplet-repeats associated with fragile X syndrome have low affinity to histone [2,6–9]. Both the pure and interrupted ATTCT pentanucleotide repeats associated with spinocerebellar ataxia type 10 have strong affinities to histones *in vitro* [1].

Friedreich's ataxia (FRDA), an autosomal recessive neurodegenerative disease, is the most common hereditary ataxia. FRDA is caused by the expansion of GAA triplet-repeats in the first intron of the frataxin (*FXN*) gene (OMIM 606829) [10,11]. Although early biochemical studies have documented that GAA triplet-repeats adopt a non-B-DNA structure, such as triplexes and 'sticky' DNA, no experimental evidence has demonstrated the presence of non-B-DNA structure in the chromosomal *FXN* in cells isolated from FRDA patients [12–16]. Recently, increasing evidence has shown that GAA triplet-repeat-induced chromatin changes may be responsible for the *FXN* mRNA deficits. Saveliev et al. found that expanded GAA triplet-repeats served as a source of HP-1-mediated heterochromatin to silence a nearby reporter gene [17]. Subsequently, several studies have suggested that the expansion of the FRDA GAA triplet-repeats can cause epigenetic changes, such as H3-K9 hypermethylation and reductions in histone acetylation. The GAA triplet-repeats can also recruit heterochromatin-binding proteins. Further, the formation of heterochromatin on intron 1 of the *FXN* gene may trigger *FXN* gene silencing through an intrinsic blocking of transcription initiation or elongation [11,18–27]. Ruan et al. found that the efficiency of nucleosome assembly on pure GAA_{22} repeat duplex was only half that of the pUC control DNA. Compared with duplex DNA *in vitro*, the GAA/GAA/TTC triplex further reduced nucleosome assembly efficiency [10].

Previous works have focused primarily on the chromatin structure of disease-length GAA triplet-repeats, and less information is available on how GAA triplet-repeats affect the local nucleosome positioning in healthy individuals. However, it is important to consider the behavior of both healthy and disease length repeat tracts at the nucleosome level.

Analyzing the nucleosome distribution around GAA triplet-repeats across an entire organism's genome is more feasible with the production of high-resolution maps of *in vivo* nucleosome positions that were generated by high-throughput sequencing techniques. In this study, the distribution of the nucleosome occupancy around GAA triplet-repeats throughout the human genome was analyzed based on published high-throughput experimental data [28]. Moreover, differences in chromatin structure among DNA fragments containing GAA triplet-repeats, CAG triplet-repeats, and Widom 601 sequences were investigated using chromatin reconstitution assay. The results suggested that GAA triplet-repeats have lower affinity to histone octamers and could change local nucleosome positioning.

Abbreviations: *FXN*, frataxin; FRDA, Friedreich's ataxia; EM, electron microscopy; AUC, analytical ultracentrifugation.

* Corresponding authors.
E-mail addresses: liguohong@sun5.ibp.ac.cn (G. Li), nmcailu@163.com (L. Cai).

http://dx.doi.org/10.1016/j.ygeno.2015.06.010
0888-7543/

2. Results

2.1. GAA triplet-repeat distribution in the human genome

To determine the frequencies of GAA triplet-repeats in the human genome, all of the uninterrupted GAA triplet-repeat tracts consisting of five or more repeated units were identified in the human genome (Fig. 1). As expected, the frequency of the GAA triplet-repeats was different for each chromosome. For example, the frequency of GAA triplet-repeats in chromosome 1 was the highest, and that in chromosome Y was the lowest. In contrast, the frequency of GAA triplet-repeat tracts decreased roughly exponentially with repeat length. The most frequent GAA triplet-repeat copies were 5, 6, 7, and 8, with more than one hundred frequencies. The frequency of the GAA triplet-repeats of 9 to 25 copies ranged from 47 to 13. In particular, only 6 fragments were found that contain more than 32 GAA triplet-repeats in the genome.

2.2. Nucleosome occupancy around GAA triplet-repeat regions

To investigate whether GAA triplet-repeats affected nucleosome arrangement *in vivo*, we first analyzed the nucleosome occupancy in the vicinity of repeat sites in the human genome. As shown in Fig. 2A, the center of the repeats showed strong depletion of nucleosomes in activated $CD4^+$ T cells, which are flanked by a small nucleosome valley. The nucleosome occupancy on approximately 200 bp downstream of GAA triplet-repeats approached the nucleosome level on sequences selected randomly from chromosomes 1–5 (Supplemental Fig. S1). This result suggested that GAA triplet-repeats affected local nucleosome positioning *in vivo*.

For comparison, the nucleosome occupancy across other repeat tracts was also analyzed. Long, pure CAG triplet-repeats are stronger nucleosome positioning elements *in vitro* [4,8]. However, no nucleosome positioning peaks were detected at the short CAG triplet-repeats *in vivo* (Fig. 2B), although the short CAG triplet-repeats were also not shown to alter nucleosome occupancy *in vitro*. Similar to GAA triplet-repeats, the centers of the TAA and CAA triplet-repeats also showed strong nucleosome depletion (Supplemental Figs. S2 and S3).

2.3. Comparison of nucleosome occupancy surrounding GAA triplet-repeats in different T cells

A recent study suggested that nucleosome location varies in different cell environments, and that several nucleosomes are either removed or shifted thus making the DNA sequence more accessible to regulatory factors [28]. To determine whether the nucleosome structure on GAA triplet-repeats was dynamically regulated in different cell environments, we compared the nucleosome distributions of the repeat regions in both activated and resting $CD4^+$ T cells. As shown in Fig. 3, no obvious differences in nucleosome profiles near GAA triplet-repeats were observed between activated and resting $CD4^+$ T cells. However, a significant increasing tendency (paired-sample t test, $P < 1.7e\text{-}59$) in nucleosome occupancy, corresponding to all of the GAA triplet-repeat tracts, was found in the resting cells compared with the activated cells.

2.4. A-tracts at the 5′ end of GAA triplet-repeats enhance the local nucleosome depletion in vivo

Surprisingly, when we analyzed the distribution of GAA triplet-repeats, A-tracts were frequently adjacent to GAA triplet-repeats in the human genome. We then tested the A contents in 10-bp upstream regions of GAA triplet-repeats, using other triplet-repeats as controls, such as TAA, CAC, CAG, CTG, and TTG. As shown in Fig. 4A, for all of 5–8 trinucleotide repeats, the percentages of GAA triplet-repeats with 7–9 and 10 adenine nucleotides in 10-bp regions in their 5′ ends were 24.2% and 33.7%, respectively. The corresponding percentages were 19% and 3.3% for TAA triplet-repeats. Almost 5% of CAC or CAG triplet-repeats exhibited more than 6 adenine nucleotides in the upstream region compared with only slightly more than 0.2% of the CTG or TTG triplet-repeats.

To test the contribution of A-tracts at the 5′ end of the repeats to the GAA triplet-repeats inducing nucleosome depletion, GAA-containing sequences were classified into two groups according to the presence of 0–6 or 7–10 adenine nucleotides in the 10-bp upstream regions of repeats. The nucleosome occupancy mapping around GAA5 and GAA6 triplet-repeats are each plotted for activated and resting T cells in Fig. 4B and Supplemental Fig. S4. Consistent with the findings mentioned above, pronounced nucleosome depletion regions were found in the vicinity of all of the triplet-repeats. A significant difference (paired-sample t test, $P < 8e\text{-}38$) was found in average nucleosome occupancy among the different A contents upstream of the triplet-repeats. Compared with valleys containing a lower A content with 0–6 adenosines, the deeper nucleosome level valleys surrounding the repeats were observed on the sequences containing 7–10 adenosines. In addition, we investigated the contribution of downstream GAA triplet-repeats to nucleosome depletion around the A-tract. The A-tract in the genome can be divided into two categories: A-tracts with and without downstream GAA triplet-repeats. All of the sequences were aligned with the A-tract start position. Supplemental Fig. S5 showed that the A-tract excluded the local nucleosome and that nucleosome depletion was positively correlated with A-tract length. The local nucleosome level on the A-tract with downstream GAA triplet-repeats was obviously lower than that on the A-tract without downstream GAA triplet-repeats. These results suggested that the GAA triplet-repeats could induce local nucleosome depletion *in vivo* and that the A-tract upstream of repeats could further enhance the depletion surrounding GAA triplet-repeats.

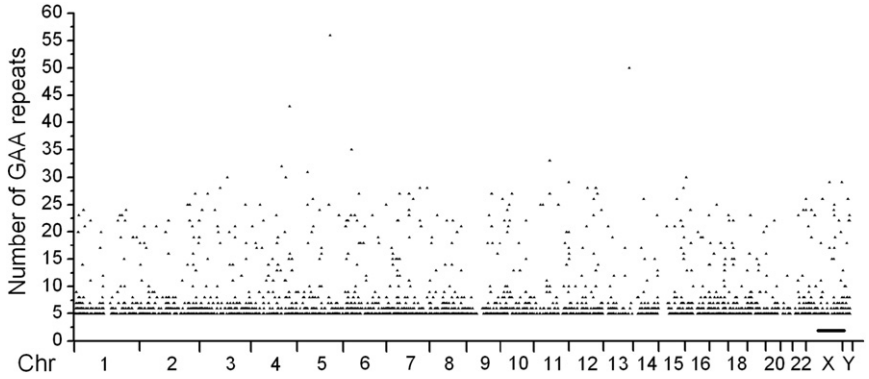


Fig. 1. Distributions of GAA trinucleotide repeats in the human genome. The y-axis denotes repeat copy numbers (repeats number ≥ 5), and the x-axis denotes chromosome length. The chromosome length bar is 1E8 bp.

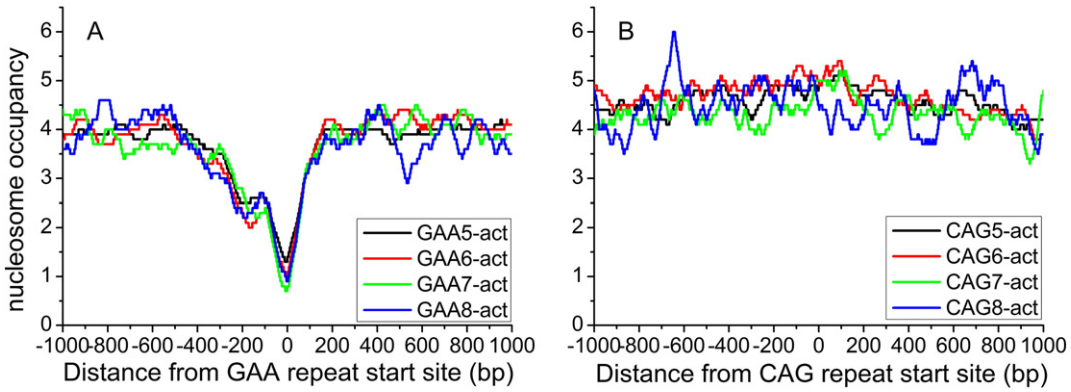


Fig. 2. Nucleosome occupancy level in a 2000-bp region relative to the start site of (A) GAA triplet-repeats and (B) CAG triplet-repeats. The different colors and styles of the curves represent the distribution of average nucleosome occupancy around the repeated tract from 5 to 8 copies. GAA5 (N = 1344), GAA6 (N = 550), GAA7 (N = 191), GAA8 (N = 108), CAG5 (N = 578), CAG6 (N = 238), CAG7 (N = 120), and CAG8 (N = 58) were evaluated, and N denotes the number of the statistical samples.

We also analyzed the nucleosome occupancy patterns on TAA and CAA triplet-repeats with a similar technique. Compared with valleys containing 0–6 adenosines, the deeper nucleosome valleys surrounding CAA triplet-repeats were observed on the sequences containing 7–10 adenosines (Supplemental Fig. S6). However, no obvious difference was found in average nucleosome occupancy between the TAA triplet-repeats with different upstream A contents (Supplemental Fig. S7). Furthermore, while the CAA and TAA triplet-repeats were adjacent to the downstream region of A-tracts, the local nucleosome depletion surrounding A-tracts was also obviously enhanced (Supplemental Figs. S8 and S9).

The effects of several triplet-repeat sequences and A-tracts on the distribution of nucleosome occupancy were analyzed statistically. Here, we preferred to pay greater attention to special regions of the genome. The first example was the region of the *FXN* gene in healthy individuals. A poly$(A)_{16}$ segment was adjacent to the 5′ end of GAA6 triplet-repeats in the first intron of the *FXN* gene. The nucleosome profiles in the region from exon 1 to part of intron 1 of the *FXN* gene for activated (top panel) and resting (middle panel) T cells are shown in Supplemental Fig. S10. Total nucleosome depletion was found from the A-tract to approximately 100 bp downstream of repeats in activated T cells. Compared with the weak nucleosome occupancy at GAA triplet-repeats, remarkable nucleosome depletion from approximately 60 bp upstream of the A-tract to the triplet-repeat start site was observed in resting T cells. These data indicated that nucleosome depletion on GAA triplet-repeats in the *FXN* genes of healthy individuals was consistent with the statistical results.

We further analyzed nucleosome occupancy on the six fragments containing more than 32 GAA triplet-repeats and their flanking sequences in the human genome. Rare A-tracts neighboring the 5′ end

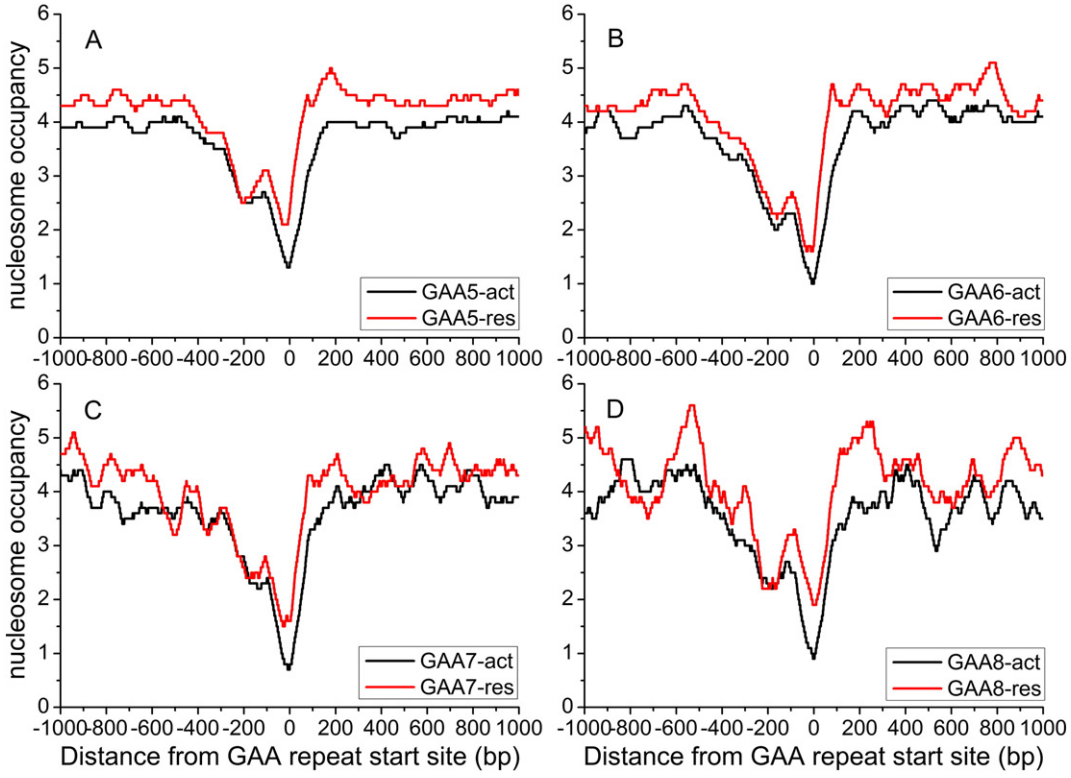


Fig. 3. Nucleosome occupancy level in a 2000-bp region relative to the start site of (A) GAA5 triplet-repeats, (B) GAA6 triplet-repeats, (C) GAA7 triplet-repeats, and (D) GAA8 triplet-repeats in activated or resting $CD4^+$ T cells.

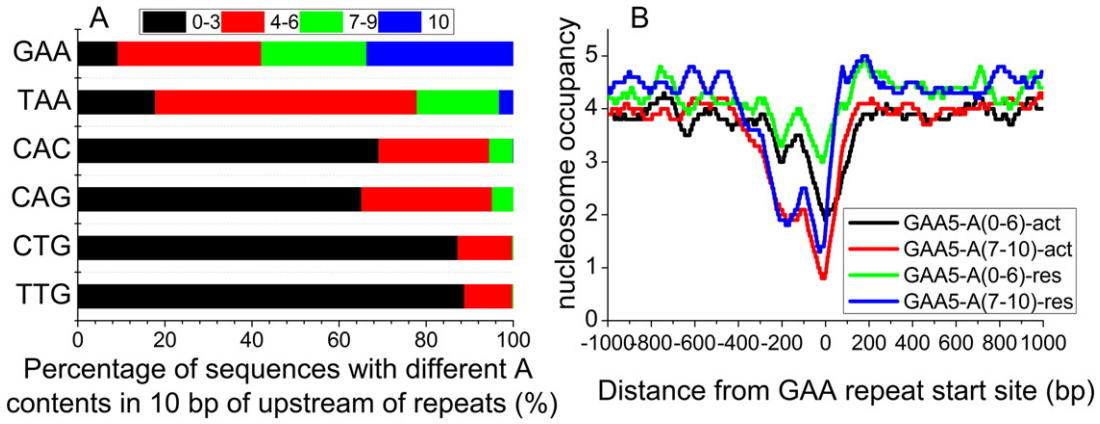


Fig. 4. Adenosines contents at the 5′ end of GAA triplet-repeats and their effect on the local nucleosome structure *in vivo*. A. A contents in a 10-bp upstream region of GAA, TAA, CAC, CAG, CTG, and TTG triplet-repeats. Different color plots in each bar denote the sequence percentages with 0–3, 4–6, 7–9, or 10 A in 10-bp upstream triplet-repeats, respectively. B. Nucleosome occupancy level on GAA5 triplet-repeat-containing templates with 0–6 or 7–10 adenosines in a 10-bp upsteam region of 5 GAA triplet-repeat start sites in activated or resting $CD4^+$ T cells. A(0–6) and A(7–10) denote the A contents in the 10-bp upstream region of GAA triplet-repeats, and act and res denote activated and resting $CD4^+$ T cells, respectively. GAA5-A(0–6) (N = 594) and GAA5-A(7–10) (N = 750) were evaluated, and N denotes the number of the statistical samples.

of GAA triplet-repeats were found (Supplemental Table S1). We computed the nucleosome occupancy over these six selected fragments in activated and resting T cells. As shown in Supplemental Fig. S11, total nucleosome depletion over GAA32, GAA33, GAA35, GAA43, and GAA56 triplet-repeats and partial nucleosome depletion on GAA50 triplet-repeats were found in activated T cells. No obvious difference was found in nucleosome distribution between activated cells and resting cells for GAAn triplet-repeats (n = 33, 35, 43, 50, and 56) (Supplemental Fig. S12).

2.5. GAA triplet-repeats change the local chromatin structure in vitro

To examine whether GAA triplet-repeats affected the local chromatin structure *in vitro*, we constructed recombinant plasmids containing GAA7, GAA27, GAA42, CAG7, CAG27, and CAG44 triplet-repeats and the Widom 601 sequence in multiple cloning sites of pUC19 (Fig. 5A). The GAA and CAG triplet-repeat lengths used were healthy-length repeat tracts. Using a recombinant histone octamer that was expressed and purified from bacteria and that lacked all posttranslational modifications (Fig. 5B), chromatins were assembled on plasmids-containing GAA and CAG triplet-repeats and the 601 sequence. The successful formation and alignment of nucleosomes were initially confirmed by micrococcal nuclease digestion (Supplemental Fig. S13).

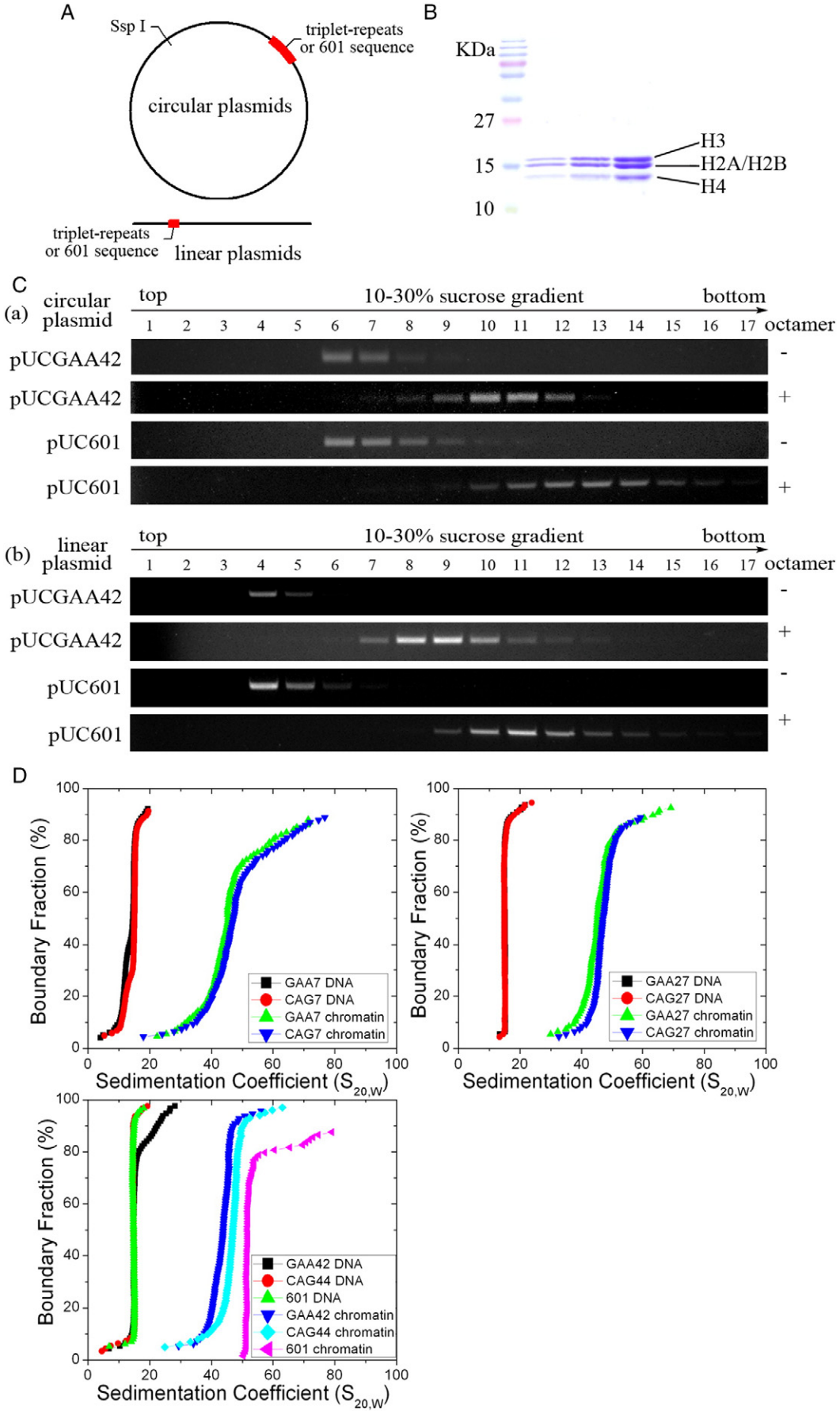
First, we examined the chromatin structure assembled on the GAA42- and 601-containing circular recombinant DNA templates using electron microscopy (EM) and sucrose density gradient centrifugation. EM images showed that the chromatins assembled with two plasmids displayed slight compaction, and no obvious differences were found in chromatin structure (Supplemental Fig. S14). To further examine the extent of chromatin assembly, the reconstituted chromatin was analyzed by sucrose gradient analysis. Whereas naked DNA alone remained in fractions 6–7 (Fig. 5C (a)), chromatin migrated toward the center or bottom of the sucrose gradient (Fig. 5C (a), fractions 9–12 of chromatin on pUCGAA42 and fractions 10–15 of chromatin on pUC601). A shift of approximately 2 fractions to the bottom was observed when the GAA42 triplet-repeat insert was replaced by the 601 sequence. These data suggested that chromatin arrays on the GAA triplet-repeat-containing template adopted a looser and more open structure than that on the 601-containing template.

The formation of one nucleosome on a closed circular DNA template could cause a change in the linking number of approximately −1.0 [29–31]. To eliminate the topological effects caused by DNA supercoils in the nucleosome assembly assay, we reconstituted chromatin on the linear plasmid templates (Fig. 5A). The reconstituted chromatins were analyzed by EM to detect any visible changes of the chromatin structure on the linear template with GAA42 triplet-repeats and 601 sequence inserts (Supplemental Fig. S14). Chromatin assembled *in vitro* displayed a “beads on a string” structure. Sucrose gradient analysis revealed that whereas linear DNA alone remained close to the top (Fig. 5C (b), fraction 4), the chromatins assembled on pUCGAA42 and on pUC601 migrated toward fractions 7–11 and fractions 9–14 of the sucrose gradient, respectively. Similar to the shift in the case of the circular DNA templates, a shift of approximately 2 fractions to the bottom was also found for bands of chromatin assembled on linear pUC601 compared with linear pUCGAA42. The results were similar to the aforementioned results of chromatin reconstitution assay on circular templates.

For the sedimentation detection of reconstituted chromatin on plasmids, the lengths of plasmids likely influence the observed differences between plasmids containing GAA42 triplet-repeats and the plasmid containing the 601 sequence. The recombinant plasmid lengths containing GAA42 triplet-repeats and the 601 sequence were 2816 and 2873 base pairs, respectively. Comparing the recombinant plasmid length, the approximately 57-bp length differences may have slightly effect on the reconstituted chromatin structure *in vitro*. To further test the effects of repeat sequences on chromatin structure quantitatively, the analytical ultracentrifugation (AUC) in sedimentation velocity experiments was analyzed for assembled chromatin with GAA-containing circular plasmids, using CAG-containing, and 601-containing plasmids as controls. The S_{ave} for the chromatin on the GAA42-containing plasmid and on the 601-containing plasmid was 43 S and 51 S, respectively, which agreed with the sucrose gradient data. The S_{ave} values of the chromatin on the GAA (n) triplet-repeats (n = 7, 27, 42) were slightly smaller than those on the CAG (n) triplet-repeats (n = 7, 27, 44) (Fig. 5D). This finding was consistent with the results derived from the previous analysis of the nucleosome occupancy on GAA and CAG triplet-repeats in the present work. The results also suggested that GAA triplet-repeat inserts destabilized the ability of recombinant plasmids to assemble nucleosomes.

3. Discussion

In the present study, the effects of GAA triplet-repeats on the local chromatin structure were investigated based on nucleosome positioning mapping data of the human genome *in vivo* and on an *in vitro* chromatin reconstitution assay. The results showed that the GAA triplet-repeats could induce local nucleosome depletion *in vivo*, and the A-tract upstream of repeats could further enhance the depletion surrounding GAA triplet-repeats. Compared with 601 sequence and CAG triplet-repeat inserts, the GAA triplet-repeat inserts in recombinant plasmids decreased nucleosome formation ability in the *in vitro*


A
Ssp I
triplet-repeats
or 601 sequence
circular plasmids
triplet-repeats
or 601 sequence
linear plasmids
B
KDa
27
15
10
H3
H2A/H2B
H4
C
(a) circular plasmid
top
10-30% sucrose gradient
bottom
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 octamer
pUCGAA42 -
pUCGAA42 +
pUC601 -
pUC601 +
(b) linear plasmid
top
10-30% sucrose gradient
bottom
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 octamer
pUCGAA42 -
pUCGAA42 +
pUC601 -
pUC601 +
D
Boundary Fraction (%)
Sedimentation Coefficient (S20,W)
GAA7 DNA
CAG7 DNA
GAA7 chromatin
CAG7 chromatin
GAA27 DNA
CAG27 DNA
GAA27 chromatin
CAG27 chromatin
GAA42 DNA
CAG44 DNA
601 DNA
GAA42 chromatin
CAG44 chromatin
601 chromatin

reconstitution assay, which indicates that the GAA triplet-repeats are weak nucleosome-positioning elements.

The weak nucleosome assembly on the GAA triplet-repeats was likely due to the abundance of AA dinucleotides. Nucleosome positioning along the genome might be determined primarily by the preference of DNA sequences [32–38] and by external factors, such as chromatin remodeling, DNA methylation, histone variants, post-translational modifications, and polymerase II binding [28,38,39]. An *in vivo* genome-wide survey of nucleosome positions and reconstituted chromatin structures *in vitro* suggested that local nucleosome positioning around GAA triplet-repeats was mediated primarily by the primary sequence. The homopolymeric stretches of deoxyadenosine nucleotides were identified as strong determinants of nucleosome organization, such as poly(dA:dT). The poly(dA:dT) tract disfavors nucleosome formation [37,40] and has been strongly associated with nucleosome depletion *in vivo* [34,41] and *in vitro* [34,39,42]. Besides, DNA flexibility could also affect nucleosome assembly [43]. Given the flexibility of dinucleotides, the AA dinucleotide is relatively stiff. Of the 12 trinucleotides, the flexibility of GAA is slightly higher than that of CAA and AAA [44]. Our previous work also indicated that the A-rich region (especially the poly-A region) is relatively rigid [45]. The weak flexible GAA trinucleotides in the sequence may have caused greater free energy for the bending of the repeats, allowing them to wrap around histones. Therefore the lots of energetic cost decrease nucleosome assembly on the triplet-repeats and the flanking sequence, which might alter the local nucleosome positioning around GAA triplet-repeats.

Our data indicated that nucleosome depletion is present on healthy-length GAA triplet-repeats in the genomes of unaffected individuals. The low-level nucleosome occupancy on GAA triplet-repeats is conducive to the transcription of the *FXN* gene.

Nucleosome positioning is dynamically regulated during the cell cycle and in response to regulation signals *in vivo*. Nucleosome arrays with high nucleosome occupancy are more feasible for forming the heterochromatin structure. However, nucleosome occupancy is not the only factor to affect the formation of the heterochromatin structure *in vivo*. At high densities of nucleosomes along the DNA, nucleosome will be positioned over lower-affinity sites, while nucleosomes will favor high-affinity sites at low densities [35,37]. In our work, the nucleosome depletion pattern on GAA triplet-repeats was investigated in cells obtained from unaffected individuals. In the disease condition, the nucleosome occupancy and dynamic repositioning may be primarily affected by epigenetic signals or other external factors rather than by the preferences of DNA sequences. The GAA triplet-repeats with low affinity to octamers may have higher nucleosome occupancy in response to some regulatory signals, although these responses have high energy demands to change their conformations.

Taken together, our results suggested that nucleosome depletion on GAA triplet-repeats was present and that GAA triplet-repeats had a weak ability to form nucleosomes in unaffected individuals. These results could answer questions regarding the chromatin structure associated with shorter-than-average expanded GAA triplet-repeats in FRDA, which is one of the definitive questions for the understanding of the pathogenesis of FRDA at the chromatin level [21]. However, several questions remain unclear. For instance, how the GAA triplet-repeat signal heterochromatin formation remains a subject of debate [12], and some strong evidence is also needed to investigate whether and how the longer expansive GAA triplet-repeats in intron 1 of the *FXN* gene affect the changes in epigenetic hallmarks in cells isolated from FRDA patients. Further study by chromatin structure analysis of GAA triplet-repeat loci in the first intron of the *FXN* gene across cellular and animal models is required to better understand the mechanisms of the *FXN* gene silencing.

4. Materials and methods

4.1. Genomic DNA and nucleosome positioning data

Whole human genome sequences (version: hg18) were downloaded from the UCSC Genome Browser (http://genome.ucsc.edu/). The human nucleosome occupancy data in activated and resting human $CD4^+$ T cells were obtained from published experimental data (http://dir.nhlbi.nih.gov/papers/lmi/epigenomes/hgtcellnucleosomes.aspx) [28]. All of the uninterrupted GAA triplet-repeat tracts consisting of five or more repeated units were identified in the human genome. Repeats lacking nucleosome occupancy information were excluded. One thousand base pair upstream and downstream of the start site of the repeats were selected. The average values of the nucleosome occupancy on all of the selected DNA fragments containing repeats were computed statistically. Customized C programs were used to examine the distributions of triplet-repeats and nucleosome levels around triplet-repeats across the entire genome.

4.2. Preparation of DNAs and recombinant histone octamer

The recombinant plasmid pUC(GAA)42 containing GAA42 was constructed as described in the Hashem-mutation document [46]. $(GAA)_7$, $(CTT)_7$, *Eco*RI adaptors (EP adaptor), and *Bam*HI adaptors (PB adaptor) were chemically synthesized (Sangon, China). $(GAA)_7$, $(CTT)_7$, and the two adaptor oligonucleotides that were phosphorylated using ATP and kinase (NEB), were precipitated, and were then resuspended in TEN (10 mM Tris–HCl, pH 7.6, 1 mM EDTA and 50 mM NaCl). Complementary pairs of oligonucleotides were hybridized by boiling and slowly cooling to 4 °C overnight. Hybridized $(GAA)_7$ / $(CTT)_7$ were ligated overnight at 4 °C in 30 μL of ligation buffer with 20 units of T4 DNA ligase (NEB). The resulting $[(GAA) \cdot (CTT)]_{42}$ polymer was then ligated at 4 °C to hybridized EP adaptor and PB adaptor. The fragments were subsequently double digested with *Eco*RI and *Bam*HI and were separated on a 5% polyacrylamide gel. The purified fragments were then ligated into the *Eco*RI and *Bam*HI sites in the multiple cloning site of pUC19. pUC(GAA)7, pUC(CAG)7, pUC(GAA)27, pUC(CAG)27, and pUC(CAG)44 were constructed using similar methods. The recombinant plasmid pUC601 was constructed by adding one copy of the 601 sequence to the *Bam*HI and *Pst*I sites of the pUC19 vector. The recombinant plasmids were prepared sufficiently. The closed plasmids were digested with *Ssp*I (NEB) to prepare the linear plasmids. The chromatins were assembled on closed or linear DNA templates.

The expression and purification of histones were performed according to methods published previously [47,48]. Each histone was expressed and purified from *Escherichia coli* BL21 cells containing pET-histone expression plasmids. The four core histones in unfolding buffer (7 M guanidinium HCl, 20 mM Tris–HCl, pH 7.5, 10 mM DTT) were mixed at equimolar ratios. The mixture was dialyzed against at least three changes of 1 L of refolding buffer (2 M NaCl, 10 mM Tris–HCl, pH 7.5, 1 mM Na-EDTA, and 5 mM 2-mercaptoethanol). The second and third dialysis steps were performed overnight. The histone octamer was maintained at 4 °C. Precipitated proteins were removed by centrifugation for 30 min at 20,000 × *g*. Samples were concentrated to a final volume of 0.5 mL before purification through a Superdex S200 filtration column (GE Healthcare). Confirmation of

Fig. 5. Chromatin assembly assay of the recombinant plasmids containing GAA triplet-repeats, CAG triplet-repeats and Widom 601 sequence *in vitro*. A. Schematic diagram of the reconstituted plasmids. GAA or CAG triplet-repeats were inserted into the *Eco*RI and *Bam*HI sites of pUC19, and the Widom 601 sequence was added into *Pst*I and *Bam*HI sites of the same vector. The closed plasmids were digested with *Ssp*I and became linear DNA. B. SDS-PAGE analysis of bacterially expressed and purified recombinant octamer. C. Sucrose gradient fractions of the assembled chromatin analyzed by agarose gel electrophoresis and ethidium bromide staining. D. Sedimentation coefficient distribution plots for GAA7-, GAG7-, GAA27-, CAG27-, GAA42-, CAG44-, and 601-containing plasmid DNA and chromatin.

the purity and stoichiometry of the histone octamer was performed using SDS-PAGE on 15% gels with Coomassie Brilliant Blue staining and the concentration was determined using an extinction coefficient at 276 nm.

4.3. Nucleosome array assembly reaction in vitro

For *in vitro* structure investigation, chromatin arrays were assembled using the salt-dialysis method as previously described [47,49]. The reconstitution reactions were conducted with histone octamer and pUC(GAA)n (n = 7, 27, 42), pUC(CAG)n (n = 7,27,44) and pUC601 DNA templates. All of the steps were performed at 4 °C. Eight micrograms of each DNA were incubated in 30 μL reactions containing 10 mM Tris–HCl (pH 8.0), 1 mM EDTA (pH 8.0), 2 M NaCl, and 9.6 μg of histone octamer (added last). The samples were placed in a microdialysis apparatus with 6–8 kDa dialysis tubing. They were then placed in a beaker containing high-salt buffer (10 mM Tris–HCl, pH 8.0, 2 M NaCl, 1 mM EDTA), which was continuously diluted by slowly pumping in TE buffer (10 mM Tris–HCl, pH 8.0, 1 mM EDTA) to a lower concentration of NaCl from 2 M to 0.6 M over a period of 16 h. After this period, the low-salt buffer beaker was replaced with HE buffer (10 mM HEPES, pH 8.0, 0.1 mM EDTA) lacking any NaCl, and the nucleosomes were dialyzed for an additional minimum of 3 h with constant stirring.

4.4. Analysis of nucleosome arrays

For each reconstitution, chromatin was analyzed by EM imaging, sucrose density gradient centrifugation (10%–30%) or sedimentation velocity with AUC.

For imaging with EM, reconstituted chromatin samples were prepared using DNA concentrations of 20 μg/mL in measurement buffer (10 mM HEPES, pH 8.0, 0.1 mM EDTA). The samples were fixed with 0.4% glutaraldehyde (Fluka) in the same buffer for 30 min on ice. Then, 2 mM spermidine was added to the sample solution to enhance the absorption of the chromatins by the glow-charged carbon-coated grids. The samples were applied to the EM grids and were incubated for 2 min and then blotted. The grids were washed stepwise with increasing amounts of ethanol/water in 20 mL baths of 0%, 25%, 50%, 75%, and 100% ethanol solution for 4 min, each at room temperature, and they were then air dried and shadowed with tungsten at an angle of rotation of 10°. The samples were examined using a FEI Tecnai G2 Spirit 120 kV transmission electron microscope [49].

For sucrose gradients, chromatin samples were loaded on a sucrose gradient in HEN buffer (10 mM HEPES, 0.2 mM EDTA, 25 mM NaCl). The sucrose gradients were centrifuged for 16 h at 24,000 rpm in a Beckman 60 Ti rotor. Fractionation was performed manually (250 μL fractions), and samples were loaded with SDS on a 1% agarose gel, stained with ethidium bromide, and destained in water.

For AUC sedimentation velocity experiments, chromatin samples containing 18 μg of DNA were prepared in measurement buffer (10 mM HEPES at pH 8.0, 0.1 mM EDTA). The sedimentation experiments were performed on a Beckman Coulter ProteomeLab XL-I using a four-hole An-60Ti rotor. Samples with an initial absorbance at 260 nm of 0.5–0.8 were equilibrated for 2 h at 20 °C under a vacuum in a centrifuge prior to sedimentation. The absorbance at 260 nm was measured in continuous scan mode during sedimentation at 30,000 × *g* for naked DNA and at 22,000 × *g* for chromatin DNA in 12-mm double-sector cells. The data were analyzed using enhanced van Holde–Weischet analysis and Ultrascan II software, version 9.9 revision 1504. The $S_{20,w}$ values (sedimentation coefficient corrected for water at 20 °C) were calculated with a partial specific volume of 0.622 mL/g for chromatin, and the buffer density and viscosity were adjusted. The average sedimentation S_{ave} coefficients were determined at the boundary midpoint [49].

Disclosure statement

The authors have no conflict of interest to declare.

Acknowledgments

This work was supported by Natural Science Foundation of China (61271448 to L. C.) and Natural Science Foundation of Inner Mongolia (2013MS0514 to H. Z. and 2014MS0312 to Y. X.).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ygeno.2015.06.010.

References

[1] K.A. Hagerman, H. Ruan, K.N. Edamura, T. Matsuura, C.E. Pearson, Y.H. Wang, The ATTCT repeats of spinocerebellar ataxia type 10 display strong nucleosome assembly which is enhanced by repeat interruptions, Gene 434 (2009) 29–34.
[2] Y.H. Wang, Chromatin structure of repeating CTG/CAG and CGG/CCG sequences in human disease, Front. Biosci. 12 (2007) 4731–4741.
[3] D.J. Mulvihill, K. Nichol Edamura, K.A. Hagerman, C.E. Pearson, Y.H. Wang, Effect of CAT or AGG interruptions and CpG methylation on nucleosome assembly upon trinucleotide repeats on spinocerebellar ataxia, type 1 and fragile X syndrome, J. Biol. Chem. 280 (2005) 4498–4503.
[4] C.B. Volle, S. Delaney, CAG/CTG repeats alter the affinity for the histone core and the positioning of DNA in the nucleosome, Biochemistry 51 (2012) 9814–9825.
[5] Y.H. Wang, J. Griffith, Expanded CTG triplet blocks from the myotonic dystrophy gene create the strongest known natural nucleosome positioning elements, Genomics 25 (1995) 570–573.
[6] Y.H. Wang, J. Griffith, Methylation of expanded CCG triplet repeat DNA from fragile X syndrome patients enhances nucleosome exclusion, J. Biol. Chem. 271 (1996) 22937–22940.
[7] J.S. Godde, S.U. Kass, M.C. Hirst, A.P. Wolffe, Nucleosome assembly on methylated CGG triplet repeats in the fragile X mental retardation gene 1 promoter, J. Biol. Chem. 271 (1996) 24325–24328.
[8] Y.H. Wang, R. Gellibolian, M. Shimizu, R.D. Wells, J. Griffith, Long CCG triplet repeat blocks exclude nucleosomes: a possible mechanism for the nature of fragile sites in chromosomes, J. Mol. Biol. 263 (1996) 511–516.
[9] S. Datta, M.P. Alam, S.S. Majumdar, A.K. Mehta, S. Maiti, N. Wadhwa, V. Brahmachari, Nucleosomal occupancy and CGG repeat expansion: a comparative analysis of triplet repeat region from mouse and human fragile X mental retardation gene 1, Chromosom. Res. 19 (2011) 445–455.
[10] H. Ruan, Y.H. Wang, Friedreich's ataxia GAA · TTC duplex and GAA · GAA · TTC triplex structures exclude nucleosome assembly, J. Mol. Biol. 383 (2008) 292–300.
[11] D. Kumari, K. Usdin, Is Friedreich ataxia an epigenetic disorder? Clin. Epigenetics 4 (2012) 2.
[12] J.M. Gottesfeld, J.R. Rusche, M. Pandolfo, Increasing frataxin gene expression with histone deacetylase inhibitors as a therapeutic approach for Friedreich's ataxia, J. Neurochem. 126 (2013) 147–154.
[13] D. Herman, K. Jenssen, R. Burnett, E. Soragni, S.L. Perlman, J.M. Gottesfeld, Histone deacetylase inhibitors reverse gene silencing in Friedreich's ataxia, Nat. Chem. Biol. 2 (2006) 551–558.
[14] N. Sakamoto, P.D. Chastain, P. Parniewski, K. Ohshima, M. Pandolfo, J.D. Griffith, R.D. Wells, Sticky DNA: self-association properties of long GAA.TTC repeats in R.R.Y triplex structures from Friedreich's ataxia, Mol. Cell (1999) 465–475.
[15] R.D. Wells, DNA triplexes and Friedreich ataxia, FASEB J. 22 (2008) 1625–1634.
[16] K. Ohshima, L. Montermini, R.D. Wells, M. Pandolfo, Inhibitory effects of expanded GAA.TTC triplet repeats from intron I of the Friedreich ataxia gene on transcription and replication *in vivo*, J. Biol. Chem. 273 (1998) 14588–14595.
[17] A. Saveliev, C. Everett, T. Sharpe, Z. Webster, R. Festenstein, DNA triplet repeats mediate heterochromatin-protein-1-sensitive variegated gene silencing, Nature 422 (2003) 909–913.
[18] D. Kumari, R.E. Biacsi, K. Usdin, Repeat expansion affects both transcription initiation and elongation in Friedreich ataxia cells, J. Biol. Chem. 286 (2011) 4209–4215.
[19] M. Baralle, T. Pastor, E. Bussani, F. Pagani, Influence of Friedreich ataxia GAA noncoding repeat expansions on pre-mRNA processing, Am. J. Hum. Genet. 83 (2008) 77–88.
[20] E. Greene, L. Mahishi, A. Entezam, D. Kumari, K. Usdin, Repeat-induced epigenetic changes in intron 1 of the frataxin gene and its consequences in Friedreich ataxia, Nucleic Acids Res. 35 (2007) 3383–3390.
[21] Y.K. Chutake, W.N. Costello, C. Lam, S.I. Bidichandani, Altered nucleosome positioning at the transcription start site and deficient transcriptional initiation in Friedreich ataxia, J. Biol. Chem. 289 (2014) 15194–15202.
[22] J.M. Gottesfeld, Small molecules affecting transcription in Friedreich ataxia, Pharmacol. Ther. 116 (2007) 236–248.
[23] S. Al-Mahdawi, R.M. Pinto, O. Ismail, D. Varshney, S. Lymperi, C. Sandi, D. Trabzuni, M. Pook, The Friedreich ataxia GAA repeat expansion mutation induces comparable epigenetic changes in human and transgenic mouse brain and heart tissues, Hum. Mol. Genet. 17 (2008) 735–746.

[24] I. Castaldo, M. Pinelli, A. Monticelli, F. Acquaviva, M. Giacchetti, A. Filla, S. Sacchetti, S. Keller, V.E. Avvedimento, L. Chiariotti, S. Cocozza, DNA methylation in intron 1 of the frataxin gene is related to GAA repeat length and age of onset in Friedreich ataxia patients, J. Med. Genet. 45 (2008) 808–812.
[25] D. Marmolino, F. Acquaviva, Friedreich's ataxia: from the (GAA)n repeat mediated silencing to new promising molecules for therapy, Cerebellum 8 (2009) 245–259.
[26] T. Punga, M. Bühler, Long intronic GAA repeats causing Friedreich ataxia impede transcription elongation, EMBO Mol. Med. 2 (2010) 120–129.
[27] M.V. Evans-Galea, N. Carrodus, S.M. Rowley, L.A. Corben, G. Tai, R. Saffery, J.C. Galati, N.C. Wong, J.M. Craig, D.R. Lynch, S.R. Regner, A.F. Brocht, S.L. Perlman, K.O. Bushara, C.M. Gomez, G.R. Wilmot, L. Li, E. Varley, M.B. Delatycki, J.P. Sarsero, *FXN* methylation predicts expression and clinical outcome in Friedreich ataxia, Ann. Neurol. 71 (2012) 487–497.
[28] D.E. Schones, K. Cui, S. Cuddapah, T.Y. Roh, A. Barski, Z. Wang, G. Wei, K. Zhao, Dynamic regulation of nucleosome positioning in the human genome, Cell 132 (2008) 887–898.
[29] A. Lusser, J.T. Kadonaga, Strategies for the reconstitution of chromatin, Nat. Methods 1 (2004) 19–26.
[30] R.T. Simpson, F. Thoma, J.M. Brubaker, Chromatin reconstituted from tandemly repeated cloned DNA fragments and core histones: a model system for study of higher order structure, Cell 42 (1985) 799–808.
[31] A. Loyola, G. LeRoy, Y.H. Wang, D. Reinberg, Reconstitution of recombinant chromatin establishes a requirement for histone-tail modifications during chromatin assembly and transcription, Genes Dev. 15 (2001) 2837–2851.
[32] Y. Bao, C.L. White, K. Luger, Nucleosome core particles containing a poly(dA.dT) sequence element exhibit a locally distorted DNA structure, J. Mol. Biol. 361 (2006) 617–624.
[33] H.E. Peckham, R.E. Thurman, Y. Fu, J.A. Stamatoyannopoulos, W.S. Noble, K. Struhl, Z. Weng, Nucleosome positioning signals in genomic DNA, Genome Res. 17 (2007) 1170–1177.
[34] N. Kaplan, I.K. Moore, Y. Fondufe-Mittendorf, A.J. Gossett, D. Tillo, Y. Field, E.M. LeProust, T.R. Hughes, J.D. Lieb, J. Widom, E. Segal, The DNA-encoded nucleosome organization of a eukaryotic genome, Nature 458 (2009) 362–366.
[35] A.E. Rapoport, Z.M. Frenkel, E.N. Trifonov, Nucleosome positioning pattern derived from oligonucleotide compositions of genomic sequences, J. Biomol. Struct. Dyn. 28 (2011) 567–574.
[36] E. Segal, Y. Fondufe-Mittendorf, L. Chen, A. Thåström, Y. Field, I.K. Moore, J.P. Wang, J. Widom, A genomic code for nucleosome positioning, Nature 442 (2006) 772–778.
[37] E. Segal, J. Widom, Poly(dA:dT) tracts: major determinants of nucleosome organization, Curr. Opin. Struct. Biol. 19 (2009) 65–71.
[38] J.Y. Wang, J. Wang, G. Liu, Calculation of nucleosomal DNA deformation energy: its implication for nucleosome positioning, Chromosom. Res. 20 (2012) 889–902.
[39] Y. Zhang, Z. Moqtaderi, B.P. Rattner, G. Euskirchen, M. Snyder, J.T. Kadonaga, X.S. Liu, K. Struhl, Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo, Nat. Struct. Mol. Biol. 16 (2009) 847–852.
[40] J.D. Anderson, J. Widom, Poly(dA-dT) promoter elements increase the equilibrium accessibility of nucleosomal DNA target sites, Mol. Cell. Biol. 21 (2001) 3830–3839.
[41] G.C. Yuan, Y.J. Liu, M.F. Dion, M.D. Slack, L.F. Wu, S.J. Altschuler, O.J. Rando, Genome-scale identification of nucleosome positions in *S. cerevisiae*, Science 309 (2005) 626–630.
[42] T. Raveh-Sadka, M. Levo, U. Shabi, B. Shany, L. Keren, M. Lotan-Pompan, D. Zeevi, E. Sharon, A. Weinberger, E. Segal, Manipulating nucleosome disfavoring sequences allows fine-tune regulation of gene expression in yeast, Nat. Genet. 44 (2012) 743–750.
[43] S. Metzenberg, On the formation of nucleosomes within the FMR1 trinucleotide repeat, Am. J. Hum. Genet. 59 (1996) 252–253.
[44] W.K. Olson, N.L. Marky, R.L. Jernigan, V.B. Zhurkin, Influence of fluctuations on DNA curvature: a comparison of flexible and static wedge models of intrinsically bent DNA, J. Mol. Biol. 232 (1993) 530–554.
[45] L. Tsai, L.F. Luo, A statistical mechanical model for predicting B-DNA curvature and flexibility, J. Theor. Biol. 207 (2000) 177–194.
[46] V.I. Hashem, W.A. Rosche, R.R. Sinden, Genetic assays for measuring rates of (CAG).(CTG) repeat instability in *Escherichia coli*, Mutat. Res. 502 (2002) 25–37.
[47] P.N. Dyer, R.S. Edayathumangalam, C.L. White, Y. Bao, S. Chakravarthy, U.M. Muthurajan, K. Luger, Reconstitution of nucleosome core particles from recombinant histones and DNA, Methods Enzymol. 375 (2004) 23–44.
[48] M. Noda, S. Uchiyama, A.R. McKay, A. Morimoto, S. Misawa, A. Yoshida, H. Shimahara, H. Takinowaki, S. Nakamura, Y. Kobayashi, S. Matsunaga, T. Ohkubo, C.V. Robinson, K. Fukui, Assembly states of the nucleosome assembly protein 1(NAP-1) revealed by sedimentation velocity and non-denaturing MS, Biochem. J. 436 (2011) 101–112.
[49] P. Chen, J. Zhao, Y. Wang, M. Wang, H. Long, D. Liang, L. Huang, Z. Wen, W. Li, X. Li, H. Feng, H. Zhao, P. Zhu, M. Li, Q.F. Wang, G. Li, H3.3 actively marks enhancers and primes gene transcription via opening higher-ordered chromatin, Genes Dev. 27 (2013) 2109–2124.